Birth weight and cognitive ability in adulthood

Citation for published version:
Grove, B, Lim, SJ, Gale, C & Shenkin, S 2017, 'Birth weight and cognitive ability in adulthood: A systematic review and meta-analysis', *Intelligence*, vol. 61, pp. 146-158. https://doi.org/10.1016/j.intell.2017.02.001

Digital Object Identifier (DOI):
10.1016/j.intell.2017.02.001

Published In:
Intelligence

Abstract

Birth weight is associated with a range of adult health outcomes. In childhood, there is a positive association between birth weight – in the normal range (>2,500 g) – and cognitive ability, but no systematic review has yet assessed this effect across adult life. We aimed to synthesise published studies assessing the relationship between birth weight and general cognitive ability in non-clinical adult populations (≥ 18 years). Nineteen studies (N = 1,122,858), mean participant age ranged from 18 to 78.4 years, fulfilled the inclusion criteria, of which eight could be included in a random-effects meta-analysis. Birth weight was associated with cognitive ability in adulthood, with each kilogram increase in birth weight associated with a 0.13 SD increase in general or fluid intelligence (95% CI [0.07, 0.19]). There was considerable heterogeneity in the effect size (I^2 = 97.8%, 95% CI [97.2, 98.4], p < .001). The association was similar after correcting for gestational age and parental social class where data were available. The effect size was larger for participants aged < 60 years than those aged 60 years or over. There is a modest association between birth weight and cognitive ability in adulthood that may diminish at older ages.

Birth Weight and Cognitive Ability in Adulthood: A Systematic Review and Meta-Analysis

Lower birth weight is associated with adverse outcomes across the lifespan. The concept of the 'Developmental Origins of Health and Disease' (Barker, 2004) has suggested that factors which influence the prenatal environment may also influence health outcomes in adult life. These include somatic outcomes such as infant mortality (Wardlaw, Blanc, Zupan, & Ahman, 2004), all-cause adult mortality (Baker, Olsen, & Sørensen, 2008), cardiovascular disease (Barker et al., 1993; Stein et al., 1996), stroke (Eriksson, Forsen, Tuomilehto, Osmond, & Barker, 2000), and type 2 diabetes (Eriksson, Forsen, Osmond, & Barker, 2003). This relationship extends to neuropsychological outcomes, where lower birth weight has been associated with outcomes such as schizophrenia (Abel et al., 2010), depression (de Mola, de França, de Avila Quevedo, & Horta, 2014; Wojcik, Lee, Colman, Hardy, & Hotopf, 2013), and cognitive ability in childhood (Shenkin, Starr, & Deary, 2004). Birth weight, especially when corrected for gestational age, is a useful marker of prenatal development, and can be influenced by placental insufficiency, maternal malnutrition, lower parental social class, genetic and epigenetic factors, and increased altitude of birth (Feil & Fraga, 2012; Jensen & Moore, 1997; Kramer, 1987).

Low birth weight (LBW < 2,500 g) babies have poorer outcomes physically and cognitively than normal birth weight (NBW) controls (Hack, Klein, & Taylor, 1995): a recent meta-analysis identified an association between LBW and poorer cognitive performance in adolescence and young adulthood, with NBW adolescents and adults scoring 7.63 IQ points higher than low birth weight participants (95% [5.95, 9.31]), reduced to 4.98 IQ points after adjusting for publication bias (95% CI [3.20, 6.77]) (Kormos, Wilkinson, Davey, & Cunningham, 2014), with the effect size reducing with increasing age. Some studies have investigated the association between birth weight in the normal range (≥ 2,500g) and cognitive ability in childhood (see review in Shenkin et al., 2004; Heinonen et al., 2008;

Lawlor et al., 2005; Lawlor et al, 2006; Yang, Lynch, Susser, & Lawlor, 2008). There is some evidence that IQ may decline at the highest birthweights (>4.5kg) (Shenkin et al., 2004). The positive association between birth weight in the normal range and cognitive ability in childhood was small: e.g. 0.81 IQ points per SD of birth weight z score adjusted for age and gender at age 5 to 6; 1.30 at age 7 to 9, and 1.44 at age 11 to 12, attenuating to 0.28, 0.67 and 0.52 points after adjusting for family characteristics (Lawlor et al, 2006; Yang, Lynch, Susser, & Lawlor, 2008). The effect is negligible at the individual level, but could have an impact at a population level, and has been a driver for assessing the impact of improving maternal health and the impact of socioeconomic influences on long term outcomes. However, observational data cannot be used to recommend interventions to increase birth weight, as there could be unintended consequences: e.g. increasing fetal weight could increase the risk of complications of labour.

Cognitive ability is generally very stable across the lifespan, in the absence of pathology. For example, Deary, Whalley, Lemmon, Crawford, and Starr (2000) identified a correlation of 0.63 between Moray House Test scores at age 11 and 77 years, and 0.73 after adjusting for the sample's ability range. Similarly, the stability coefficient for general intelligence in a cohort of Swedish men was 0.95 between 18 and 50 years, and 0.86 between 18 and 65 years (Rönnlund, Sundström, & Nilsson, 2015). As suggested by the authors, this stability fits well with the parieto-frontal integration theory (P-FIT; Jung & Haier, 2007), which links cognitive stability with neural stability, and cognitive decline to decreased neural stability in old age. Factors from early life can persist into old age, although it has been debated whether this is due to permanent programming in early life or an ageing-related accumulation of deficits (e.g. Kirkwood & Melov, 2011, Walker, 2011). Proponents of permanent programming theories stress that even small differences in early life conditions can influence later health outcomes (Gavrilov, Leonid, Natalia, & Gavrilova, 2004). The stability of cognitive ability across the lifespan highlights the importance of determining the relationship between early-life factors and cognitive ability in adulthood.

No systematic review has yet assessed the relationship between birth weight and cognitive ability across the entirety of adulthood and across the entire range of birth weight, to assess if the association found in childhood persists, strengthens or weakens.

We aimed to conduct a systematic review and meta-analysis on studies that assessed the relationship between birth weight across the normal range and performance on any cognitive assessment in a nonclinical adult population.

Methods

Protocol and Registration

We registered the protocol for this review with the International Prospective Register of Systematic Reviews (PROSPERO) prior to the formal search. Permanent link: http://dx.doi.org/10.15124/CRD42015020380.

Eligibility Criteria

Eligible studies assessed adult participants ($M_{age} \geq 18$ years) of normal birth weight on at least one cognitive test. We considered all observational study types for inclusion. We excluded studies if participants were members of, or matched controls for, a clinical population or a LBW group (< 2,500 g). We also excluded studies where cognitive ability was only assessed by a measure of cognitive success (e.g. education, employment). We did not limit publications on language or publication date. Studies in which a standardised beta coefficient was provided for the relationship between birth weight and a measure of fluid or general intelligence were included in the meta-analysis. If this was not published in the paper, we contacted the study author.

Identification of Studies

Information sources. We ran an electronic search via OvidSP in EMBASE, PsycINFO and Medline (including in-process and non-indexed citations) in September 2015. We conducted a forward citation search on all studies identified for inclusion in the systematic review, and checked the reference lists of included studies for any further relevant articles.

Search. The search was devised with an experienced librarian, and adapted for each database. Briefly: titles, abstracts and subject headings were searched for terms relating to birth weight AND cognition. (Supplement 1). Animal studies, and studies only including children, were excluded from the search.

Study selection. One reviewer (BJG) screened all titles and abstracts against the eligibility criteria. A second reviewer (JSL) independently reviewed a subset of these studies. Any areas of uncertainty resolved via discussion with CRG or SDS. When studies were from the same cohort we planned to use the paper with the most comprehensive (and recent) data.

Data Extraction

The data extraction form was based on the Cochrane Consumers and Communication Review Group's template (CCCRG, 2009), and revised following piloting (Supplement 2). One reviewer (BJG) conducted data extraction, and a second reviewer (YCH) checked all data extracted. Any disagreements were resolved after discussion with CRG or SDS. If the paper did not contain the relevant analysis for inclusion in the meta-analysis, but included a fluid or general cognitive measure, we attempted to contact the corresponding author for further information. We requested the standardised beta coefficient for the relationship between birth weight (per kilogram increase) and standardized fluid ability score, unadjusted, adjusted for gestational age (where possible) and adjusted for both gestational age and a measure of socioeconomic status at birth (where possible).

Risk of bias. Risk of bias was assessed by use of an adapted version of the Quality in Prognostic Studies (QUIPS) tool (Hayden, van der Windt, Cartwright, Côté, & Bombadier, 2013) and was conducted by one reviewer (BJG).

Results

Meta-analysis. Where data were available, we used meta-analysis, conducted with STATA version 13 (StataCorp 2013), to obtain an overall estimate for the effect and to quantify the estimate's uncertainty. A meta-analysis was conducted for the crude association between birth weight (per kilogram increase) and standardized fluid cognitive ability score. We used DerSimonian and Laird random effect models to calculate the pooled effect for each cohort, which accounts for between-sample variation (Deeks, Altman, & Bradburn, 2001). We examined the heterogeneity of the estimates between studies using the I^2 statistic (with 95% confidence intervals). This statistic quantifies the percentage of total variation across studies due to heterogeneity rather than chance. An I^2 statistic of 25%, 50% or 75% suggests low, moderate or high heterogeneity respectively (Higgins, Thompson, Deeks, & Altman, 2003). We produced forest plots for the overall unadjusted effect. We also examined, through additional meta-analyses, how the effect would change when correcting for gestational age and both gestational age and socioeconomic status at birth. Finally, we conducted subgroup meta-analyses to quantify the effect for different participant age brackets. We assessed risk of publication bias through a funnel plot of studies included in the meta-analysis.

Studies not included in meta-analysis. For studies that did not provide the data required for inclusion in the meta-analysis, but contained relevant information, the results were described in more detail. As we aimed to assess how the effect might change over different stages of life, we presented studies in order of increasing participant age, stratified into different age brackets.

Results

Nineteen studies were included in the systematic review (*N* = 1,122,858) (Table 1). Mean participant age ranged from 18 to 78.4 years. Eight of these studies were also included in the meta-analysis.

Study selection

Of 8,899 unique citations (Figure 1), a second reviewer (JSL) screened a subset of 1,240 studies (13.9%), identifying eight as eligible for inclusion. There was full agreement on these eight. The full-texts of 101 potentially relevant studies were retrieved and screened against the inclusion and exclusion criteria by one reviewer. Two papers were identified from forward citation searching. No more studies were identified through reference lists.

Repeated data. Nine papers were excluded for assessing the same cohort as a separate paper included in the systematic review. We chose the study with the largest amount of data relevant to the review: (1) Bergvall, Iliadou, Tuvemo and Cnattingius (2006b) (and therefore excluded Bergvall, Iliadou, Johansson, Tuvemo, & Cnattingus, 2006a; Gunnell, Harrison, Rasmussen, Fouskakis, & Tynelius, 2002; Lundberg et al., 2010; Lundgren, Cnattingius, Jonsson, & Tuvemo, 2001; Lundgren, Cnattingius, Jonsson, & Tuvemo, 2003; Yang, Bergvall, Cnattingius, & Kramer, 2010); (2) Kristensen et al. (2014), excluding Eide, Øyen, Skjærven, & Bjerkedal, 2007; and Eriksen, Sundet, & Tambs, 2010; (3) Richards, Hardy, Kuh, & Wadsworth (2001) was included as it reported more cognitive assessments than Richards, Hardy, Kuh, & Wadsworth (2002). No excluded study reported conflicting results to the included studies.

Figure 1 about here

Study selection for meta-analysis. Of 19 papers considered for the meta-analysis, six did not use a relevant cognitive measure. One paper included the required information for meta-analysis (Skogen et al 2013) and 12 included relevant cognitive measures but did not have the required information. We contacted the 11 corresponding authors (two papers had the same corresponding author), and received the relevant data for seven of the 12 papers. Eight papers were included in the unadjusted meta-analysis (Dawes et al., 2015: De Rooij et al., 2010: Kristensen et al., 2013: Martyn et al., 1996: Raikkonen et al., 2009: Shenkin et al., 2004: Skogen et al., 2013; Victora et al., 2015), and five of these in the adjusted analysis (De Rooij et al., 2010: Kristensen et al, 2014; Raikkonen et al, 2009: Shenkin et al., 2004: Victora et a l., 2015). We also contacted the corresponding author on a potentially eligible paper (cognitive ability measured against ponderal index) (Zhang et al., 2009), but did not receive a reply.

Table 1 about here

Characteristics of Included Studies

Fourteen studies were from Europe, three from the United States, one from Australia and one from Brazil (Table 1). Birth weight was determined by self-report in three studies and participants were retrospectively matched to their birth records in all others. Thirteen studies reported on multiple different cognitive tests and 11 studies provided a general or fluid intelligence score (five of which also reported on multiple tests) (Table 1). Cognitive decline/change was reported in four studies, two of which determined decline by comparing tests of crystallised and fluid intelligence, and two were longitudinal re-test comparisons. Ten studies were rated as medium for risk of bias, seven as low and two as high (Supplement 3). Table 2 shows the results of individual studies, including potential confounders adjusted for and risk of bias ratings. Where categorical results were provided by birth weight group the outcomes are summarised in Table 2 and full results reported for unadjusted and adjusted results respectively are in Supplements 4 and 5.

Table 2 about here

Studies Included in Meta-Analysis

Participants. One of the studies was from Brazil (Victora et al., 2015), and the other seven were from European countries. Two studies were limited to male participants only (Kristensen et al., 2014; Räikkönen et al., 2009). One study had separate data available for five different cohorts (Martyn, Gale, Sayer, & Fall, 1996). Mean participant age was between 18 years and 78 years.

Cognitive assessment. Two studies used information from IQ tests given at military conscription (Kristensen et al., 2014; Räikkönen et al., 2009). Two studies reported general IQ scores (Dawes et al., 2015; Victora et al., 2015). Two studies used the Alice-Heim Test (fourth version) (de Rooij, Wouters, Yonker, Painter., & Roseboon, 2010; Martyn et al., 1996). Two studies used composite scores from several cognitive subtests (Shenkin, Deary, & Starr, 2009; Skogen, Øverland, Smith., & Mykletun, 2013).

Risk of bias. Risk of bias scores were medium for five of the studies (Dawes et al., 2014; de Rooij et al., 2010; Martyn et al., 1996; Skogen et al., 2013; Victora et al., 2015). In each case, this was due to either no key confounders being included in the adjusted model, or no adjusted model being provided. Two of these studies provided both adjusted conditions requested (de Rooij et al., 2010; Victora et al., 2015). In each case, reasons for medium risk of bias were either resolved or irrelevant to the unadjusted analysis. The remaining three studies were low risk of bias (Kristensen et al, 2014; Räikkönen et al., 2009; Shenkin et al., 2009).

Publication bias. A funnel plot showed some asymmetry, with a lack of smaller studies showing no beneficial effect, indicating that some publication bias may exist (Supplement 6).

Meta-Analysis Results

Figure 2 presents the forest plot for the difference in standardized fluid cognition score per kilogram increase of birth weight across the normal range. Results are ordered by ascending participant age. Higher birth weight was associated with increased cognitive ability (z = .15, 95% CI [0.09, 0.22]). There was considerable heterogeneity in the effect size (I^2 = 91.7%, 95% CI [97.2, 98.4], p < .001).

The effect was similar after adjusting for gestational age (z = 0.15, 95% CI [0.07, 0.24], I^2 = 80.7%, 95% CI [54.8, 91.8], p < .001) and both gestational age and parental social class at birth (z = 0.14, 95% CI [0.06, 0.22], I^2 = 80.9%, 95% CI [55.4, 91.8], p < .001) (Supplement 7-8)

Examination of the forest plots suggested that the effect of birth weight declined with age. We made a post-hoc decision to dichotomise by study participants' mean age: < 60 years or ≥ 60 years, where the effect appeared to diminish. Birth weight was associated with cognitive ability for those < 60 years, with considerable heterogeneity (z = .15, 95% CI [0.08, 0.23], I^2 = 94.9%, 95% CI [91.8, 96.9], p < .001), and not associated with cognitive ability in those >=60 years, with less heterogeneity (z = 0.07, 95% CI [-0.02, 0.16], I^2 = 34.6%, p < .001), though confidence intervals around this latter I^2 statistic were wide so the true degree of heterogeneity is uncertain.

Studies not Included in Meta-Analysis

Young adulthood (mean age 18 to 39 years)

Individual studies. Pearce, Mann, Singh, & Sayers (2014) studied 283 18 year olds born to self-identifying Australian Aboriginal mothers. Mean birth weight was 3,000 g, which is lower than all other studies where mean birth weight was reported. Participants were tested on three measures from the CogState battery (simple and choice reaction time, and working memory). Simple reaction time was notably high (median = 345.74 ms, IQR

[283.35, 468.52]). Birth weight related to simple reaction time, but not choice reaction time or working memory. This was consistent both before and after adjusting for gestational age, residential status and participant age. After adjustments, simple reaction time was faster by 76.39 ms for each kilogram increase in birth weight for gestational age (95% CI [24.43, 128.16], $p = .004$). A medium risk of bias rating was given as loss to follow-up information was unclear, alongside a small sample size. Note that this is the only study to include reaction time measures, and that the mean SRT values are much higher than Western populations at similar ages. The fact that there was a statistically significant association between birth weight and SRT, but not CRT, as well as genetic and cultural differences, mean the results should be treated with caution.

Bergvall et al. (2006b) reported on 356,206 Swedish men who were given a general intelligence test at military conscription. Results were standardised to stanine scores ($M = 5$, $SD = 2$). After adjustments, the risk of low intellectual performance (scores of ≤ 2) increased for those whose birth weight for gestational age was more than two standard deviations below the mean (OR = 1.22, 95% CI [1.13, 1.33]). High birth weight (> 2 *SDs*) was not related to risk of low intelligence, indicating a non-linear relationship (OR = 0.98, 95% CI [0.90, 1.06]). Results were similar before and after adjustment for key confounders. One study of the same cohort which was not included (Gunnell et al., 2002) analysed the relationship between birth weight and unadjusted performance on each of the four separate subtests, identifying positive linear relationships in each, without testing for non-linear relationships. Risk of bias was low.

Sørensen et al. (1997) reported on 4,300 Danish men tested at military conscription. Birth weight was positively associated with IQ score. Before adjustments, mean score on the Boerge Priens test ranged from 39.9 ($SD = 9.3$) in participants ≤ 2500 g, increasing to 44.6

(*SD* = 9.5) for participants with birthweights > 4500 g. A quadratic spline regression was fitted after adjusting for socioeconomic factors at birth, mother's age, and several birth parameters (gestational age, length at birth, parity), showing that the trend becomes negative above 4,200 g. Summary effect size was not reported. Risk of bias was low.

Richards et al. (2001) reported on 3,115 participants from the 1946 British birth cohort, tested on reading comprehension at age 26 years. Both before and after adjusting for parental social factors, birth order and sex, reading comprehension was significantly positively associated with birth weight. Participants with birth weight ≤ 2500g scored significantly lower than the reference group (z = -0.33, 95% CI [-0.51, -0.16]). Although the overall trend was significant (p = 0.001), this was largely due to the difference between the low birth weight and reference group. Adjustment for birth order was reported to increase the coefficient for the highest birth weight group. The medium risk of bias rating arose from a lack of correction for gestational age.

Flensborg-Madsen and Mortensen (2015) reported on 937 participants from the Copenhagen Perinatal Cohort. The association was determined via a t-test between the scores of participants under 3,300 g and above 3,300 g. Birth weight was not significantly associated with IQ score, although this relationship was marginal (p = 0.06). A medium risk of bias rating was given as the association was not adjusted for any confounders. This paper met the eligibility criteria as both groups were within the range of normal birth weight, however, it does not provide an accurate measure of how the relationship changes across the entire range of birth weight (as the birth weight cut off is 3,300 g rather than 2,500 g).

Summary. In five studies including seven cognitive tests in young adulthood, there was a small but statistically significant association between birth weight and cognitive ability. Statistical significance was identified in two of three general IQ tests and tests for simple reaction time and reading comprehension. Significance was not identified for one IQ test or tests of choice reaction time and working memory.

Middle age (mean age 40 - 60 years).

Individual Studies. Richards et al. (2001), reported on the longitudinal follow-up of 2,575 participants tested again at 43 years of age (note tests performed at age 26 on same cohort reported above). Different tests were used, so longitudinal comparisons cannot be easily made. Birth weight was not significantly associated with verbal memory ($p = 0.08$), search speed ($p = 0.79$) or search accuracy ($p = 0.78$). Results were consistent after adjusting for the same factors as previously discussed. Risk of bias was moderate due to lack of correction for gestational age.

Factor-Litvak et al. (2011) reported from the United States Early Determinants of Adult Health study, with the 474 participants stratified by two test centres and sex for analysis. After adjusting for confounders, birth weight was significantly associated with attention ($B = 0.03$, $SE = 0.012$, $p = .03$) and verbal fluency ($B = -0.1$, $SE = 0.044$, $p = .04$) in men from one test centre, but not for the women or for participants from the other test centre. Immediate recall was significantly associated with birth weight for women from the other test centre ($B = -0.048$, $SE = 0.02$, $p = .02$), but not for any other group. There were no significant association for delayed recall. A high risk of bias rating was given as no unadjusted associations were reported and no summary statistic was provided for the entire cohort; positive associations may have been false positives.

Costa et al. (2011) reported on 3,292 participants from the United States Atherosclerosis Risk in Communities cohort who had recalled their exact birth weight (17.6% of the total cohort). After adjusting for confounders (not including gestational age), a 100 g increase in birth weight was significantly related to an increase of 0.75 words in a word fluency test (95% CI [0.17, 1.33], $p = .004$). Birth weight was not significantly related to verbal recall or the Digit Symbol test. A high risk of bias rating was given as no adjusted associations were provided, the study ran the risk of exclusion bias and controlling for too many confounders

Summary. In three studies including nine cognitive tests in middle age, birth weight was not reliably associated with cognitive ability. Unadjusted results were given from one study, where birth weight was not significantly associated with verbal memory, or a search task (accuracy and speed). These were also non-significant after adjustments, alongside delayed word recall and the Digit Symbol test. There were mixed findings for tests of attention, immediate recall, delayed recall and verbal fluency, discussed below.

Older age (mean age > 60 years).

Individual Studies.

Paile-Hyvärinen et al. (2009) assessed 1,243 participants with mean age of 63.9 years, from the Helsinki Birth Cohort on five measures from the CogState battery. After adjustments, reaction time in the divided attention task was faster by 3.8% for each kilogram increase in birth weight for gestational age (95% CI [-6.5, -1.1], p = .005). Errors made in the associate learning task also decreased by 1.5% for each kilogram increase in birth weight for gestational age (95% CI [-0.1, -2.9], p = .04). Risk of bias was low.

Räikkönen et al. (2013) reported on a longitudinal follow-up of Finnish men at age 67.9 years previously tested at military conscription (Räikkönen et al., 2009). Adjustments were made for birth parameters (gestational age and parity), history of breastfeeding, adult health status, education, and socioeconomic parameters at birth. After adjustments, each standard deviation increase in birth weight for gestational age was associated with an increase of 1.31 points on a general intelligence measure (M = 20.1, SD = 4.2, 95% CI [0.06, 2.55], p = .04). This was also significant before adjustments. It is not clear the extent to which this participant group overlaps with Paile-Hyvärinen et al. (2009), although participants in this study were exclusively male, while 56.9% of participants in Paile-Hyvärinen et al.'s study were female. Risk of bias was low.

Erickson, Kritz-Silverstein, Wingard, & Barrett-Connor (2010) reported on 292 women from the United States Rancho Bernardo Study. From six cognitive tests, with 12 results reported, birth weight adjusted for age and education was related to scores on only one test, the Mini-Mental State Examination (MMSE) subtest of serial sevens (β = 0.08, p = .04). This may be a false positive. Both the serial sevens and world backwards subtests from the MMSE were reported for all participants, although they are typically used interchangeably; the world backwards test was non-significant (β = -0.00, p = .89), so this result may be a false positive. A medium risk of bias was given as a result of there being no reported unadjusted correlation, a lack of key confounders and the risk of a false positive result.

Muller et al. (2014) reported on 1,254 participants from the Icelandic Gene/Environment Susceptibility Cohort. After adjusting for sex, age and education, composite scores for memory, processing speed and executive function (calculated through eight cognitive tests) were not significantly associated with birth weight. No unadjusted association was reported, and the association was not adjusted for gestational age, resulting in a medium risk of bias rating.

Summary. There were 20 cognitive tests across four studies. Birth weight was associated with one test of full-scale IQ, both before and after adjusting for confounders. Birth weight was also significantly associated with score on the Serial 7's test, divided attention, and associate learning (hit rate). Birth weight was not significantly associated with the Buschke-Fuld Selective Reminding task (total score, long term memory, short term memory), Heaton visual copying (score, long term memory, short term memory), the Mini-mental state exam (MMSE) total score, the world backwards test from the MMSE, the Trail-Making Test B, Blessed information-memory-concentration test, category fluency, memory, processing speed, executive function, associate learning (reaction time), simple reaction time, choice reaction time, or working memory (hit rate and reaction time).

Discussion

Nineteen published studies of the association between birth weight in the normal range and general cognitive ability were identified, eight of which were included in a meta-analysis. There was a modest association between birth weight and cognitive ability in adulthood, which was independent of confounding factors like socioeconomic status at birth and gestational age. On a standard IQ scale (M = 100, SD = 15) this translates to roughly a 1.98 point increase for every kilogram increase of birth weight. While individual studies tended to not show statistical significance in older age, the meta-analysis suggests that the effect does persist beyond young adulthood through to old age, but may diminish after around 60 years of age. Separate meta-analyses according to whether the mean age of participants was < 60 or 60 years or over showed that the effect size was larger in the younger age group. It is possible that the effect diminishes with age, but remains detectible in larger-scale studies. This is supported by the fact that in the only longitudinal study to re-test participants, there was a statistically significant association between birthweight and cognitive ability in 931 men aged 61-71 years (Räikkönen et al., 2013). This is the only study to date to find a significant effect of birthweight on cognitive ability in people aged 60 or over.

Birth Weight and Cognitive Ability

The role of different factors, e.g. genetic or environmental influences, contributing to a relationship between birth weight and cognitive ability may differ with increasing age. Even in childhood, the relationship between birth weight and cognitive ability is small, with birth weight explaining about 1% of variance in cognition (Shenkin et al., 2004). Variations in birth weight can, however, have a large impact at a population level if they are associated with later outcomes. The factors which contribute to cognition in childhood are not identical to those that affect cognitive performance in adulthood; although cognitive ability tends to remain stable, genes have a greater contribution to cognitive ability in adulthood than in childhood, where the environment plays a more important role (Briley & Tucker-Drob, 2015). Both genetics and the environment also mediate cognitive change between childhood and older age (Deary et al., 2012). A study of the Danish National Birth Cohort (n = 1,782) showed that in five-year-old children, IQ, parental education and maternal IQ explained 17% of the variance, whereas birth weight explained <1% of the variance (Eriksen et al, 2013). The increasing role of environmental factors has been shown in a study relating several other perinatal measures to cognitive ability in an elderly Chinese cohort, where associations were significant until several environmental covariates were introduced to the model (Zhang et al., 2009). Our systematic review and meta-analysis suggests that the relationship between birth weight and cognition may persist beyond the fourth decade of life.

Given the environmental factors that contribute both to birth weight and cognitive development, it is possible that there would be cultural differences in any association found between the two. The studies in this systematic review were generally fairly homogeneous in terms of culture, although there was one study from a developing country (Victora et al., 2015); in this Brazilian cohort there was a significant association between birth weight and cognitive ability at a mean age of 30.2 years. As this association was not adjusted for confounders, it would be worthwhile for future research to investigate factors mediating any relationships between the perinatal environment and cognitive outcomes in developing countries, or specific racial (genetic) or ethnic groups in developed countries. The one included study (not eligible for the meta-analysis) of self-identified Aboriginal people in Australia (Pearce et al., 2014), measured reaction time (SRT and CRT) and working memory (CogState battery) developed for the assessment of diverse groups. This is the only included study which used reaction time measures, and the mean SRT values are much higher than

Western populations at similar ages (median 346ms), which would indicate a lower score on conventional IQ tests. The fact that there was a statistically significant association between birth weight and SRT, but not CRT, means the results should be treated with caution. There was no association between birth weight and working memory. Results from this sample cannot be extrapolated to other populations. This cohort differs from others in the review genetically, as well educationally, culturally and socially differences, and the results suggest that the relationship between test scores and birth weight may differ between different groups. The impact of genetics, education, cultural and social factors on the relationship between birth weight and IQ across the life course should be explored in future studies.

Five studies were male-only cohorts which all identified a positive relationship, including the only statistically significant association in older age (Bergvall et al., 2006b; Kristensen et al., 2014; Raikksonen et al., 2009; Räikkönen et al., 2013; Sørensen et al., 1997). While this may also have resulted from chance, males are more susceptible to birth insults and prenatal brain damage than females (e.g. Nunez & McCarthy, 2003) and this may have played a part in these positive associations. Future research could assess sex-specific cognitive outcomes of normal variations in birth weight.

Another aspect worth consideration is whether there is a cohort or historical effect, or variations between countries. For example, the incidence of infant mortality in Finland is currently around 0.2% (World Bank, 2015), yet the infant mortality rate in Helsinki at the inception of the Helsinki Birth Cohort (1934 - 44) ranged from around 5 to 9% (OSF, 2010). This suggests that the neurological profile of modern children may differ from many of those included in the studies in this systematic review. However, the improvements in healthcare that have resulted in decreased infant mortality might also mean that modern healthcare is more equipped to improve the outcomes of low birth weight, an idea supported by Hansen and Greisen (2004), who identified that an increase in survival of very low birth weight babies did not result in impaired intelligence at 5 years. It is noted that several studies identified poorer cognitive outcomes among the heaviest babies (Erikson et al., 2010; Kristensen et al., 2014; Richards et al., 2001; Sørensen et al., 1997); this is an important group for future study.

Strengths and Limitations of Included Studies

There are a range of potential sources of bias across included studies. Associations were largely limited to developed countries and to male participants. It is unclear whether the findings of this systematic review and meta-analysis are generalizable to developing countries and whether the effect is sex-specific.

Studies were limited to participants who were given a cognitive test and whose birth weight was available. These participant samples may not be representative of the general population, particularly in older cohorts where hospital births were less common.

Attrition bias may have also affected results. Many studies did not go into detail about loss to follow up. Better education appears to delay cognitive ageing (Meng & D’Arcy, 2012), and a bias towards educated participants remaining in studies may have influenced findings towards non-significant results. As the fetal environment is associated with several other health outcomes, those whose cognitive outcomes were most affected by the prenatal environment could have also had higher dropout rates as a result of health-related outcomes which may not have emerged until adult life.

A wide range of cognitive assessments were used for a different cognitive domains. While this allows for an overview of both general intelligence and the specific components of general intelligence, it also makes comparisons between individual studies difficult. Of the studies which included a general intelligence factor and were included in the meta-analysis, two constructed these from a variety of sub-tests (Shenkin et al., 2009; Skogen et al., 2013), and only two used the same cognitive assessment (de Rooij et al., 2010; Martyn et al., 1996). This range means that it is difficult to pinpoint whether different cognitive domains are affected by this association to differing degrees.

Studies varied in the covariates included in adjusted models. Two studies provided no adjusted condition in the published paper, as birth weight and cognitive ability was not the primary association, but did provide these upon request for use in the meta-analysis (de Rooij et al., 2010; Victora et al., 2015), and one study had no adjusted condition and was not included in the meta-analysis (Flensborg-Madsen and Mortensen, 2015). Seven more - four in older age (Costa et al., 2010; Dawes et al., 2015; Erikson et al., 2010; Martyn et al., 1996; Muller et al., 2014; Richards et al., 2001; Skogen et al., 2013) - did not adjust for gestational age as a result of lack of information available, which lessens the specificity of birth weight as a measure of fetal growth. It is possible that this could have contributed to the lack of significance in some studies. Most studies also did not provide justifications for the majority of adjustments. Furthermore, one study (Costa et al., 2011) adjusted for a total of 21 different measures (not including gestational age), which makes it difficult to assess how far participants represent the general population. Seven studies also did not provide any unadjusted information, making it difficult to assess the role of covariates in the reported effect.

Six studies also adjusted for the participant's highest educational achievement in adulthood, mostly as a proxy for socioeconomic status. While adult educational achievement is somewhat predicted by social circumstances at birth, there is substantial evidence of a relationship between intellectual ability and social mobility (upwards and downwards) as well as educational achievement (Deary et al., 2005; Strenze, 2007., Waller, 1982). If there is a causal relationship between the prenatal environment and later cognitive ability, a large part of adult educational achievement lies on the same causal pathway, as educational achievement may be a proxy for cognitive ability. Future research in this area should consider the implications of adjusting for such factors.

Participant exclusions may also introduce bias. Although most studies excluded a minimal number of participants, one study excluded 36.6% (n = 3,921) of participants (Costa et al., 2011). It is debateable whether such extensive exclusions allow for generalisability to the wider population.

Included studies in middle and older age tended to report a large number of effects. Although this allows for a comprehensive variety of cognitive assessments, any significant effects are more likely to result from chance. As a result, some significant findings run the risk of being type I errors.

The meta-analysis was also limited to studies which provided an estimate for the association between birth weight and general or fluid cognitive ability, and whose authors were able to provide additional information.

Strengths and Limitations of the Review Process

This review largely did not deviate from the protocol which was constructed prior to the formal search. This is with the exception of a change in tool used for the risk of bias assessment, and also in the decision not to include the analysis in Costa et al. (2011) where birth weight category was recalled by participants.

The forward citation search was conducted systematically, and combined with reference list searching, all current relevant studies appear to have been included. There were no restrictions on publication language, and full-texts of all potentially relevant articles were retrieved. The search identified more studies than expected, and participants were dispersed across a wide range of different ages. Despite confidence in the search strategy, the grey literature was not systematically searched, although it was scoped for relevant studies to no avail. This may have resulted in some relevant analyses not being included in this review. Furthermore, it is possible that there are unpublished studies that were not available.

There was considerable heterogeneity across studies, and this is both a strength and a weakness. As this review was specifically aimed towards exploring the relationship between birth weight and cognitive ability across different age brackets and cognitive assessments, this heterogeneity is welcomed in the context of the review's objectives. However, the nature of included studies means that it is not possible to engage in an assessment of how the relationship might manifest across different cognitive domains. There was significant heterogeneity in effect sizes from the articles included in the meta-analysis (I^2 >75%), which was expected, given the range of different factors known to contribute to both birth weight and cognitive ability, the different cognitive tests with their own scoring systems, and the range of demographics across each study. We accounted for this by the use of a random effects model, which does not assume consistency in the effect across studies.

Although the underlying implications of birth weight relate to the intrauterine environment in general, it is also noted that birth weight is not the only useful measure of the prenatal environment, even though it is the most readily available. Especially in older age, the included studies identified sporadic positive associations between cognitive ability and other perinatal measures such as biparietal diameter, head circumference, birth length, gestational age and ponderal index, all of which may relate to specific aspects of fetal growth and specific periods of gestation. For example, one of the included studies (Muller et al., 2014) identified that birth weight was positively associated with larger adult head size and brain volume in older women, which in turn were both associated with increased scores on tests of processing speed and executive function. There are also studies which assess the relationship between perinatal factors other than birth weight and cognitive ability or decline in older age (Gale, Walton & Martyn, 2003, Zhang et al., 2009). While birth weight itself is not a reliable predictor of cognitive ability or decline beyond young adulthood, it would be erroneous to conclude from the included studies that no birth parameters are related to cognitive ability or decline in middle or older age. Future reviews may consider adult cognitive outcomes of measures other than birth weight. While the aim of assessing this association is to draw conclusions about the relationship between the prenatal environment and later cognitive outcomes, such definitive conclusions cannot be drawn from birth weight data alone.

Conclusion

Birth weight is modestly association with cognitive ability in adulthood. This association is reliably found in young adulthood, but the persistence of the effect through to older age in less clear. Factors contributing to the effect may differ across different cultures and sexes. Future research should focus on specific prenatal factors and take a considered approach to covariates.

Acknowledgements

This study was not externally funded. Thanks to Sheila Fisken, Liaison Librarian, University of Edinburgh for advice on the search strategy, and to Yu-Chi Huang, Department of Psychology, University of Edinburgh for assistance in data extraction. CRG and SDS are members of The University of Edinburgh Centre for Cognitive Ageing and Cognitive Epidemiology, part of the cross council Lifelong Health and Wellbeing Initiative (MR/K026992/1). Funding from the BBSRC and MRC is gratefully acknowledged. The authors thank the following contributors for providing additional analyses of their data: Piers Dawes, Susanne De Rooij, Bernardo Horta and Christian Loret de Mola, Petter Kristensen, and Katri Räikkönen.

References

* Studies marked with an asterisk were included in the systematic review

Abel, K. M., Wicks, S., Susser, E. S., Dalman, C., Pedersen, M. G., Mortensen, P. B., & Webb, R. T. (2010). Birth weight, schizophrenia, and adult mental disorder: is risk confined to the smallest babies?. *Archives of General Psychiatry*, *67*(9), 923-930.

Pink, B., & Allbon, P. (2008). *The health and welfare of Australia's Aboriginal and Torres Strait Islander peoples*. Canberra: Commonwealth of Australia.

Baker, J. L., Olsen, L. W., & Sørensen, T. I. (2008). Weight at birth and all-cause mortality in adulthood. *Epidemiology, 19*(2), 197-203.

Barker, D. J., Godfrey, K. M., Gluckman, P. D., Harding, J. E., Owens, J. A., & Robinson, J. S. (1993). Fetal nutrition and cardiovascular disease in adult life. *The Lancet*, *341* (8850), 938-941.

Barker, D. J. P. (2004). The developmental origins of adult disease. *Journal of Epidemiology and Community Health, 58*(2), 114-115.

Bergvall, N., Iliadou, A., Johansson, S., Tuvemo, T., & Cnattingius, S. (2006a). Risks for low intellectual performance related to being born small for gestational age are modified by gestational age. *Pediatrics, 117*(3), e460-e467.

*Bergvall, N., Iliadou, A., Tuvemo, T., & Cnattingius, S. (2006b). Birth characteristics and risk of low intellectual performance in early adulthood: are the associations confounded by socioeconomic factors in adolescence or familial effects?. *Pediatrics, 117*(3), 714-721.

Briley, D. A., & Tucker-Drob, E. M. (2015). Comparing the Developmental Genetics of Cognition and Personality over the Life Span. *Journal of Personality*. Advance online publication. doi: 10.1111/jopy.12186

Cochrane Consumers and Communication Review Group (CCCRG). (2009). Data Extraction Template. Available from: http://cccrg.cochrane.org/author-resources

*Costa, A. J. L., Kale, P. L., Luiz, R. R., De Moraes, S. A., Mosley, T. H., & Szklo, M. (2011). Association between birthweight and cognitive function in middle age: The Atherosclerosis Risk in Communities Study. *Annals of Epidemiology, 21*(11), 851-856.

Davis, E. P., & Pfaff, D. (2014). Sexually dimorphic responses to early adversity: implications for affective problems and autism spectrum disorder. *Psychoneuroendocrinology, 49*, 11-25.

*Dawes, P., Cruickshanks, K. J., Moore, D. R., Fortnum, H., Edmondson-Jones, M., McCormack, A., & Munro, K. J. (2015). The Effect of Prenatal and Childhood Development on Hearing, Vision and Cognition in Adulthood. *PloS ONE*, *10*(8), e0136590.

Deary, I. J., Taylor, M. D., Hart, C. L., Wilson, V., Smith, G. D., Blane, D., & Starr, J. M. (2005). Intergenerational social mobility and mid-life status attainment: influences of childhood intelligence, childhood social factors, and education. *Intelligence, 33*(5), 455-472.

Deary, I. J., Yang, J., Davies, G., Harris, S. E., Tenesa, A., Liewald, D., ... & Visscher, P. M. (2012). Genetic contributions to stability and change in intelligence from childhood to old age. *Nature*, *482*(7384), 212-215.

Deary, I. J., Whalley, L. J., Lemmon, H., Crawford, J. R., & Starr, J. M. (2000). The stability of individual differences in mental ability from childhood to old age: follow-up of the 1932 Scottish Mental Survey. *Intelligence, 28*(1), 49-55.

Deeks, J. J., Altman, D. G., & Bradburn, M. J. (2008). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. *Systematic Reviews in Health Care: Meta-Analysis in Context* (2), 285-312.

De Mola, C. L., de França, G. V. A., de Avila Quevedo, L., & Horta, B. L. (2014). Low birth weight, preterm birth and small for gestational age association with adult depression: systematic review and meta-analysis. *The British Journal of Psychiatry, 205*(5), 340-347.

*De Rooij, S. R., Wouters, H., Yonker, J. E., Painter, R. C., & Roseboom, T. J. (2010). Prenatal undernutrition and cognitive function in late adulthood. *Proceedings of the National Academy of Sciences, 107*(39), 16881-16886.

Eide, M. G., Øyen, N., Skjærven, R., & Bjerkedal, T. (2007). Associations of birth size, gestational age, and adult size with intellectual performance: evidence from a cohort of Norwegian men. *Pediatric Research,62*(5), 636-642.

Eriksen, H. L. F., Kesmodel, U. S., Underbjerg, M., Kilburn, T. R., Bertrand, J., & Mortensen, E. L. (2013). Predictors of intelligence at the age of 5: family, pregnancy and birth characteristics, postnatal influences, and postnatal growth. PLoS One, 8(11), e79200.

Eriksen, W., Sundet, J. M., & Tambs, K. (2010). Birth weight standardized to gestational age and intelligence in young adulthood: a register-based birth cohort study of male siblings. *American Journal of Epidemiology, 172*(5), 530-536

*Erickson, K., Kritz-Silverstein, D., Wingard, D. L., & Barrett-Connor, E. (2010). Birth weight and cognitive performance in older women: the Rancho Bernardo study. *Archives of Women's Mental Health, 13*(2), 141-146.

Eriksson, J. G., Forsen, T. J., Osmond, C., & Barker, D. J. (2003). Pathways of infant and childhood growth that lead to type 2 diabetes. *Diabetes care*, *26*(11), 3006-3010.

Eriksson, J. G., Forsen, T., Tuomilehto, J., Osmond, C., & Barker, D. J. P. (2000). Early growth, adult income, and risk of stroke. *Stroke, 31*(4), 869-874.

*Factor-Litvak, P., Straka, N., Cherkerzian, S., Richards, M., Liu, X., Sher, A., ... & Goldstein, J. (2011). Associations between birth weight, preeclampsia and cognitive functions in middle-aged adults. *Journal of Developmental Origins of Health and Disease, 2*(06), 365-374.

Feil, R., & Fraga, M. F. (2012). Epigenetics and the environment: emerging patterns and implications. *Nature Reviews Genetics*, 13(2), 97-109.

*Flensborg-Madsen, T., & Mortensen, E. L. (2015). Infant developmental milestones and adult intelligence: A 34-year follow-up. *Early Human Development*, *91*(7), 393-400.

Gale CR, Walton S, Martyn CN. (2003). Foetal and postnatal head growth and risk of cognitive decline in old age. *Brain 126*(10):2273–2278.

Gunnell, D., Harrison, G., Rasmussen, F., Fouskakis, D., & Tynelius, P. E. R. (2002). Associations between premorbid intellectual performance, early-life exposures and early-onset schizophrenia Cohort study. *The British Journal of Psychiatry, 181*(4), 298-305.

Hack, M., Klein, N. K., & Taylor, H. G. (1995). Long-term developmental outcomes of low birth weight infants. *The Future of Children, 5*(1) 176-196.

Hansen, B. M., & G. Greisen (2004). Is improved survival of very-low-birthweight infants in the 1980s and 1990s associated with increasing intellectual deficit in surviving children?. *Developmental Medicine and Child Neurology 46*(12): 812-815.

Harder, T., Rodekamp, E., Schellong, K., Dudenhausen, J. W., & Plagemann, A. (2007). Birth weight and subsequent risk of type 2 diabetes: a meta-analysis. *American Journal of Epidemiology*, *165*(8), 849-857.

Hayden JA, van der Windt DA, Cartwright JL, Côté P, Bombardier C. (2013). Assessing bias in studies of prognostic factors. *Ann Intern Med. 158*(4):280-6.

Heinonen, K., Räikkönen, K., Pesonen, A. K., Kajantie, E., Andersson, S., Eriksson, J. G., ... & Lano, A. (2008). Prenatal and postnatal growth and cognitive abilities at 56 months of age: a longitudinal study of infants born at term. *Pediatrics*, *121*(5), e1325-e1333.

Higgins JP, Thompson SG, Deeks JJ, Altman DG (2003). Measuring inconsistency in meta-analyses. British Medical Journal 327, 557–560.

Jensen, G. M., & Moore, L. G. (1997). The effect of high altitude and other risk factors on birthweight: independent or interactive effects?. *American Journal of Public Health*, *87*(6), 1003-1007.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behavioral and Brain Sciences, 30*(02), 135-154.

Kirkwood, T. B., & Melov, S. (2011). On the programmed/non-programmed nature of ageing within the life history. *Current Biology, 21*(18), 701-707.

Kormos, C. E., Wilkinson, A. J., Davey, C. J., & Cunningham, A. J. (2014). Low birth weight and intelligence in adolescence and early adulthood: a meta-analysis. *Journal of Public Health, 36*(2), 213-224.

Kramer, M. S. (1987). Determinants of low birth weight: methodological assessment and meta-analysis. *Bulletin of the World Health Organization, 65*(5), 663-737.

*Kristensen, P., Susser, E., Irgens, L. M., Mehlum, I. S., Corbett, K., & Bjerkedal, T. (2014). The association of high birth weight with intelligence in young adulthood: a cohort study of male siblings. *American Journal of Epidemiology, 180*(9), 876-884.

Lawlor, D. A., Batty, G. D., Morton, S. M., Deary, I. J., Macintyre, S., Ronalds, G., & Leon, D. A. (2005). Early life predictors of childhood intelligence: evidence from the Aberdeen children of the 1950s study. *Journal of Epidemiology and Community Health*, *59*(8), 656-663.

Lawlor, D. A., Najman, J. M., Batty, G. D., O'Callaghan, M. J., Williams, G. M., & Bor, W. (2006). Early life predictors of childhood intelligence: findings from the Mater-University study of pregnancy and its outcomes. *Paediatric and Perinatal Epidemiology*, *20*(2), 148-162.

Lundberg, F., Cnattingius, S., D'Onofrio, B., Altman, D., Lambe, M., Hultman, C., & Iliadou, A. (2010). Maternal smoking during pregnancy and intellectual performance in young adult Swedish male offspring. *Paediatric and Perinatal Epidemiology, 24*(1), 79-87.

Lundgren, E. M., Cnattingius, S., Jonsson, B., & Tuvemo, T. (2001). Intellectual and psychological performance in males born small for gestational age with and without catch-up growth. *Pediatric Research, 50*(1), 91-96.

Lundgren, E. M., Cnattingius, S., Jonsson, B., & Tuvemo, T. (2003). Birth characteristics and different dimensions of intellectual performance in young males: a nationwide population-based study. *Acta Paediatrica*, *92*(10), 1138-1143.

*Martyn, C. N., Gale, C. R., Sayer, A. A., & Fall, C. (1996). Growth in utero and cognitive function in adult life: follow up study of people born between 1920 and 1943. *BMJ, 312*(7043), 1393-1396.

Meng, X., & D'Arcy, C. (2012). Education and dementia in the context of the cognitive reserve hypothesis: a systematic review with meta-analyses and qualitative analyses. *PLoS ONE, 7*(6), e3826.

*Muller, M., Sigurdsson, S., Kjartansson, O., Jonsson, P. V., Garcia, M., von Bonsdorff, M. B., ... & Launer, L. J. (2014). Birth size and brain function 75 years later. *Pediatrics, 134*(4), 761-770.

Nunez, J. L., & McCarthy, M. M. (2003). Sex differences and hormonal effects in a model of preterm infant brain injury. *Annals of the New York Academy of Sciences, 1008*(1), 281-284.

Official Statistics of Finland (OSF) (2010). Causes of death [e-publication]. Infant mortality in 1936 to 2010 . Helsinki: Statistics Finland. Retrieved from http://www.stat.fi/til/ksyyt/2010/ksyyt_2010_2011-12-16_kat_007_en.html

*Paile-Hyvärinen, M., Räikkönen, K., Kajantie, E., Darby, D., Ylihärsilä, H., Salonen, M. K., ... & Eriksson, J. G. (2009). Impact of glucose metabolism and birth size on cognitive performance in elderly subjects. *Diabetes Research and Clinical Practice,83*(3), 379-386.

*Pearce, M. S., Mann, K. D., Singh, G., & Sayers, S. M. (2014). Birth weight and cognitive function in early adulthood: the Australian Aboriginal birth cohort study. *Journal of Developmental Origins of Health and Disease, 5*(03), 240-247.

*Räikkönen, K., Forsén, T., Henriksson, M., Kajantie, E., Heinonen, K., Pesonen, A. K., ... & Eriksson, J. G. (2009). Growth Trajectories and Intellectual Abilities in Young Adulthood The Helsinki Birth Cohort Study. *American Journal of Epidemiology,170*(4), 447-455.

*Räikkönen, K., Kajantie, E., Pesonen, A. K., Heinonen, K., Alastalo, H., Leskinen, J. T., ... & Eriksson, J. G. (2013). Early life origins cognitive decline: findings in elderly men in the Helsinki Birth Cohort Study. *PLoS ONE, 8*(1), e54707.

*Richards, M., Hardy, R., Kuh, D., & Wadsworth, M. E. (2001). Birth weight and cognitive function in the British 1946 birth cohort: longitudinal population based study. *BMJ,322*(7280), 199-203.

Richards, M., Hardy, R., Kuh, D., & Wadsworth, M. E. (2002). Birthweight, postnatal growth and cognitive function in a national UK birth cohort. *International Journal of Epidemiology,31*(2), 342-348.

Rönnlund, M., Sundström, A., & Nilsson, L. G. (2015). Interindividual differences in general cognitive ability from age 18 to age 65years are extremely stable and strongly associated with working memory capacity. *Intelligence*, *53*, 59-64.

*Shenkin, S. D., Deary, I. J., & Starr, J. M. (2009). Birth parameters and cognitive ability in older age: A follow-up study of people born 1921–1926. *Gerontology, 55*(1), 92-98.

Shenkin, S. D., Starr, J. M., & Deary, I. J. (2004). Birth weight and cognitive ability in childhood: a systematic review. *Psychological bulletin, 130*(6), 989-1013.

*Skogen, J. C., Øverland, S., Smith, A. D., Mykletun, A., & Stewart, R. (2013). The impact of early life factors on cognitive function in old age: The Hordaland Health Study (HUSK). *BMC Psychology, 1*(1), 16.

*Sørensen, H. T., Sabroe, S., Olsen, J., Rothman, K. J., Gillman, M. W., & Fischer, P. (1997). Birth weight and cognitive function in young adult life: historical cohort study. *BMJ, 315*(7105), 401-403.

StataCorp (2013). Stata release 13; Statistical Software; College Station, TX: StataCorp LP.

Stein, C. E., Fall, C. H. D., Kumaran, K., Osmond, C., Barker, D. J. P., & Cox, V. (1996). Fetal growth and coronary heart disease in South India. *The Lancet*, *348*(9037), 1269-1273.

Strenze, T. (2007). Intelligence and socioeconomic success: A meta-analytic review of longitudinal research. *Intelligence, 35*(5), 401-426.

*Victora, C. G., Horta, B. L., de Mola, C. L., Quevedo, L., Pinheiro, R. T., Gigante, D. P., ... & Barros, F. C. (2015). Association between breastfeeding and intelligence, educational attainment, and income at 30 years of age: a prospective birth cohort study from Brazil. *The Lancet Global Health*, *3*(4), e199-e205.

Walker, R. F. (2011). Developmental theory of aging revisited: focus on causal and mechanistic links between development and senescence. *Rejuvenation research*, *14*(4), 429-436.

Waller, J. H. (1982). Achievement and social mobility: Relationships among IQ score, education, and occupation in two generations. *Social Biology, 29*(3-4), 255-263.

Wardlaw T, Blanc A, Zupan J & Ahman E (2004). *Low Birth Weight: Country, Regional and Global Estimates.* United Nation's Children's Fund (Unicef) and World Health Organisation (WHO), New York.

Wojcik, W., Lee, W., Colman, I., Hardy, R., & Hotopf, M. (2013). Foetal origins of depression? A systematic review and meta-analysis of low birth weight and later depression. *Psychological Medicine, 43*(01), 1-12.

World Bank (2015). Motality rate, infant (per 1,000 live births). Retrieved from http://data.worldbank.org/indicator/SP.DYN.IMRT.IN

Yang, S., Bergvall, N., Cnattingius, S., & Kramer, M. S. (2010). Gestational age differences in health and development among young Swedish men born at term. *International Journal of Epidemiology, 39*(5), 1240-1249.

Yang, S., Lynch, J., Susser, E. S., & Lawlor, D. A. (2008). Birth weight and cognitive ability in childhood among siblings and nonsiblings. *Pediatrics*, *122*(2), e350-e358.

Zhang, Z. X., Plassman, B. L., Xu, Q., Zahner, G. E. P., Wu, B., Gai, M. Y., ... & Xu, T. (2009). Lifespan influences on mid-to late-life cognitive function in a Chinese birth cohort. *Neurology, 73*(3), 186-194.

Table 1

Characteristics of Included Studies, Ordered by Participant Age

Study		Participants							Assessment
Reference	Setting	*n*	Birth year	Test year	BW (g) mean (SD)	% male	Participant age (years) range	Participant age (years) mean (SD)	Cognitive domain
Pearce et al. (2014)	Australia: Australian Aboriginal Birth Cohort	283	1987 - 90	2005 - 08	3000 (600)	46	17 - 19	18 (1.09)	Simple Reaction Time; Choice Reaction Time; Working Memory (CogState).
Bergvall et al. (2006b)	Sweden: Military Conscripts	356,206	1973 - 81	1991 - 2000	NR	100	17 - 19 (range for 99%)	NR	IQ (Swedish Conscripts Intelligence Test)
Kristensen et al. (2014)*	Norway: Military Conscripts	217,746	1967 - 76	1984 - 2003	3595 (490)	100	18 - 19 (range for 98%)	NR	IQ (Norwegian Armed Forces Draft Examination)
Sørensen et al. (1997)	Denmark: Military Conscripts	4,300	1973 - 75	1993 - 94	3471 (490)	100	18 - 20[a]	NR	IQ (Boerge Priens)
Räikkönen et al. (2009)*	Finland: Military Conscripts (from HBCS)	2,786	1934 - 44	1952 - 72	3500 (500)	100	17 - 28	20.1 (1.4)	Verbal reasoning; Visuospatial reasoning; Arithmetic reasoning (Finnish Defence Forces Draft Examination)
Richards et al. (2001)[b]	United Kingdom: NSHD	3,115	1946	1972	NR	NR	NR	26	Reading comprehension (Watts-Vernon)
Flensborg-Madsen and Mortensen (2015)	Denmark: Copenhagen Perinatal Cohort	937	1959 - 61	1982 - 74	3269 (555)	51	20 - 34	27.6 (4.3)	IQ (WAIS)
Victora et al. (2015)*	Brazil: Pelotas birth cohort	3,493	1982	2012 - 13	3225 (525)	48	NR	30.2	IQ (WAIS)
Dawes et al. (2015)*	United Kingdom: UK Biobank	503,325	NR	2006 -10	NR	45	NR	56	Processing speed (based on ‘snap’); IQ (fluid intelligence test)
Richards et al. (2001)[b]	United Kingdom: NSHD	2,575	1946	1989	NR	NR	NR	43	Verbal memory (word list learning); Search speed and accuracy (timed letter search)
Factor-Litvak et al. (2011)	United States: Early Determinants of Adult Health Study	474	NR	NR	NR	45	NR	43.8 (2.3)	Attention (SCPT); Immediate recall (WAIS Digit Symbol); Delayed recall (California verbal learning); Verbal fluency (F-A-S test)
De Rooij et al. (2010)*	Netherlands: Dutch Famine Birth Cohort	737	1943 - 47	2002 - 04	3365 (464)	47	59	NR	RT and score (AH4); RT and score (Stroop-like task); Memory retrieval and immediate recall (paragraph recall); Errors and rounds ad errors per rounds (mirror drawing)
Costa et al. (2011)	United States: ARCS	3,292	1923 - 44	1990 - 98	3500 (700)	44	51 - 70	59.2 (5.6)	Word recall (Delayed word recall); Digit Symbol (WAIS); Verbal Fluency (MAE word fluency)
Martyn et al. (1996)*	United Kingdom: Recruited Participants	1,576	1920 - 43	NR	3382 (524)	NR	48 - 74	60.9 (2.1)	AH4 score (AH4); ;Mill Hill vocabulary test.

Table 1 (*continued*)

Study		Participants							Assessment
Reference	Setting	*n*	Birth year	Test year	BW (g) mean (SD)	% male	participant age (years) range	participant age (years) mean (SD)	Cognitive domain
Paile-Hyvärinen et al. (2009)	Finland: Helsinki Birth Cohort Study	1,243	1933 - 34	2001 - 04	3354 (481)	46	NR	63.9 (2.8)	Divided attention; associate learning HR and RT; simple reaction time; choice reaction time; working memory HR and RT (CogState)
Räikkönen et al. (2013)[d]	Finland: Helsinki Birth Cohort Study	931	1934 - 44	2009	3482 (475)	100	64 - 71	67.9 (2.5)	IQ (Finnish Defence Forces Draft Examination).
Erickson et al. (2010)	United States: Rancho Bernardo Study	292	NR	1988 - 91	3357 (862)	0	55 - 89	71.1 (8.6)	Buschke total recall and LTM and STM (BFSR); Heaton copying and LTM and STM (Heaton visual reproduction); MMSE total score and serial sevens and world backwards (MMSE); Blessed (IMC); Trails B (Trail-making task B); Category fluency test.
Skogen et al. (2013)*	Norway: The Hordaland Health Study	346	1925 - 27	1997 - 99	3470 (530)	46	72 - 74	72.3	Episodic memory (Kendrick Object Learning); Trails A (Trail-making task A); Digit Symbol (WAIS); Block design (WAIS); MMSE (MMSE reduced); Verbal fluency (COWAT); IQ (composite score of all previous tests, standardised)
Muller et al. (2014)	Iceand: Age, Gene/Environment Susceptibility Cohort	1,254	1907 - 35	2002 - 06	3700 (500)	43	NR	75 (5)	Memory (California Verbal Learning); Processing speed (Figure comparison test and WAIS Digit Symbol and Stroop 1 and Stroop 2 all combined); Executive function (CaNTAB Spatial Working Memory and Digits backwards test and Stroop 3 all combined
Shenkin et al. (2009)*	United Kingdom: Recruited Participants	128	1921 - 26	2000 - 02	3330 (460)	30	75 - 81	78.4 (1.4)	Executive function (COWAT); Verbal reasoning (Moray House Test No. 12); Non-verbal reasoning (RSPM); Logical memory (WAIS Logical Memory); *g* (General intelligence composite score of all previous); Crystallised intelligence (NART).

Studies marked with * were included in the meta-analysis.

[a] Range unclear

[b] The study reported on two different age groups and has been split accordingly

[c] Follow-up an average of four years later

[d] Longitudinal subset of Räikkönen et al. (2009)

Abbreviations: AH4, Alice-Heim Test (fourth version); BSFR, Busche-Fuld Selective Reminding Task; BW, birth weight; CaNTAB, Cambridge Neuropsychological Test Automated Battery; COWAT; Controlled Oral Word Association Test; HBCS, the Helsinki Birth Cohort Study; HR, hit rate; IMC; information-memory-concentration; LTM, long-term memory; MAE, Multilingual Aphasia Examination; NART, National Adult Reading Test; MMSE, Mini-Mental State Examination; NHSD, Medical Research Council National Survey of Health and Development; RSPM, Raven's Standard Progressive Matrices; RT, reaction time; SCPT; Seidman Continuous Performance Test; SD, standard deviation; STM, short-term memory; WAIS, Wechsler Adult Intelligence Scale.

Table 2

Results of Individual Studies, Ordered by Participant Age

Study		Unadjusted Correlation			Adjusted Correlation			Confounders	Risk of Bias
Reference	Test	Coefficient	SE	p	Coefficient	SE	p		
Pearce et al. (2014)	Simple Reaction Time	*B* = –61.01	19.91	.002*	*B* = –76.39	26.41	.004*	**B**: GA; **AD**: Residential Status; **M**: Age	Medium
	Choice Reaction Time	*B* = –24.03	17.77	.13	*B* = –16.31	23.12	.48		
	Working Memory	*B* = –56.81	95.64	.55	*B* = –142.31	126.35	.26		
Bergvall et al. (2006b)	IQ (Swedish Conscripts)	†		+*	†		+*	**B**: GA, HC, length, parity **AD**: BMI, postnatal growth; year of test **M**: Age **P**: social class, education, family structure	Low
Kristensen et al. (2014)	IQ (Norwegian Armed Forces)	†		+*	†		+*	**B**: GA, birth year, birth order, parity; **M**: Age, marital status; **F**: Age, income; **P**: Education	Low
Sørensen et al. (1997)	IQ (Boerge Priens)	†		+*	†		+*	**B**: GA, length, parity **AD:**[a] **M**: Age, marital status, employment	Low
Räikkönen et al. (2009)	Verbal reasoning	NR	NR	NR			.27	**B**: GA, Parity **CH:** history of breastfeeding	Low
	Visuospatial reasoning	NR	NR	NR	*β* = .06	0.023	.008*	**AD**: Age, year of testing	
	Arithmetic reasoning	NR	NR	NR	*β* = .05	0.026	.03*	**M**: Age, height; **F**: Occupation	
Richards et al. (2001)[b]	Reading comprehension	†		.001*	†		<.001*	**B:** Birth order, sex **M:** Age, education **F:** Social class	Medium
Flensborg-Madsen and Mortensen (2015)	IQ	†		.06	NR		NR	NA	Medium
Victora et al. (2015)	IQ	†		<.001*	NR		NR	NA	Medium
Dawes et al. (2015)	IQ	*β* = 0.03		<.01*	*β* = 0.01		<.01*	**B:** Sex **AD:** Age, cardiovascular disease, diabetes,	Medium
	Reaction Time	*β* = -0.01		<.01*	*β* = -0.03		<.01*	education, smoking, socioeconomic status,	
	IQ four year change	*β* = -0.03		ns	*β* = -0.02		ns	hypertension, cholesterol **M:** Smoking	
	Reaction time four year change	*β* = 0.01		ns	*β* = 0.00		ns		
Richards et al. (2001)[b]	Verbal memory	†		.08	†		ns	**B:** Birth order, sex	Medium
	Search accuracy	†		.80	†		ns	**M:** Age, education	
	Search speed	†		.78	†		ns	**F:** Social class	
Factor-Litvak et al. (2011)	Attention			NR	*B* = -0.002	0.009	.82	**B:** GA, race	High
				NR	*B* = 0.03	0.012	.03*	**AD:** Age, smoking, socioeconomic status, sibships	
				NR	*B* = -0.017	0.011	.12		
				NR	*B* = 0.001	0.014	.96		
	Immediate Recall			NR	*B* = -0.01	0.018	.60		
				NR	*B* = 0.01	0.019	.62		
				NR	*B* = -0.048	0.02	.02*		
				NR	*B* = 0.019	0.024	.44		

Table 2 (*continued*)

Study: Reference	Study: Test	Unadjusted Correlation: Coefficient	Unadjusted Correlation: SE	Unadjusted Correlation: p	Adjusted Correlation: Coefficient	Adjusted Correlation: SE	Adjusted Correlation: p	Confounders	Risk of Bias
Factor-Litvak et al. (2011) (*Continued*)	Delayed Free Recall			NR	B = 0.011	0.009	.24		High
				NR	B = 0.008	0.01	.46		
				NR	B = -0.017	0.012	.16		
				NR	B = 0.002	0.011	.85		
	Verbal fluency			NR	B = -0.012	0.036	.74		
				NR	B = -0.1	0.044	.04*		
				NR	B = 0.055	0.042	.20		
				NR	B = 0.017	0.044	.71		
de Rooij et al. (2010)	AH4 RT	ρ = .03		ns	NR	NR	NR	NA	Medium
	AH4 score	ρ = .06		ns	NR	NR	NR		
	Stroop RT	ρ = –.01		ns	NR	NR	NR		
	Stroop score	ρ = .03		ns	NR	NR	NR		
	Memory retrieval	ρ = .01		ns	NR	NR	NR		
	Immediate recall	ρ = –.02		ns	NR	NR	NR		
	Mirror errors	ρ = –.07		ns	NR	NR	NR		
	Mirror rounds	ρ = .06		ns	NR	NR	NR		
	Mirror errors/round	ρ = –.08		ns	NR	NR	NR		
Costa et al. (2011)	Word Fluency	NR	NR	NR	B = 0.752	0.30	.004	**B:** Race, Sex, **AD:** Age, Alc, BMI, CLM, Db, education, FC, HDL, Hd, HS, Hy, LDL, MS, SI, smoking, VE; **P:** Education, Hst, Hhd, Hdb	High
	Delayed Word Recall	NR	NR	NR	B = 0.028	0.03	ns		
	Digit Symbol	NR	NR	NR	B = –0.067	0.25	ns		
Martyn et al. (1996)	AH4	†	NR	NR			.17	**B:** Social class **AD:** Age, individual dataset	Medium
	Decline	†	NR	NR			.42		
Paile-Hyvärinen et al. (2009)	Divided Attention			NR	B = –3.8	1.38	.005	**B:** GA, Sex **AD:** Age, Education (history of heart disease, depression and self-reported health status also considered but not included in adjusted model)	Low
	Associate Learning HR			NR	B = –1.5	0.71	.04		
	Associate learning RT			NR		ns	ns		
	Simple RT			NR		ns	ns		
	Choice RT			NR		ns	ns		
	Working Memory HR			NR		ns	ns		
	Working Memory RT			NR		ns	ns		
Räikkönen et al. (2014)	IQ (Finnish Defence Forces)	B = 1.04	0.51	.04*	B = 1.31	0.64	.04*	**B:** GA, parity; **CH**: breastfeeding; **AD:** Education, Hhd, Hst **M:** Age, height; **F:** Social class	Low
	Decline	r = .07	0.04	.04*	r = .08	0.04	.06		

Table 2 (*continued*)

Study: Reference	Study: Test	Unadjusted Correlation: coef	SE	p	Adjusted Correlation: coef	SE	p	Confounders	Risk of Bias
Erikson et al. (2010)	Buschke total			NR	$\beta = -0.08$		.77	**AD:** Age, education	Medium
	Buschke LTM			NR	$\beta = -0.08$		.83		
	Buschke STM			NR	$\beta = 0.00$		.97		
	Heaton visual copying			NR	$\beta = 0.05$		.63		
	Heaton visual LTM			NR	$\beta = -0.00$		.99		
	Heaton visual STM			NR	$\beta = 0.07$		.22		
	MMSE total			NR	$\beta = 0.03$		.57		
	Serial 7's			NR	$\beta = 0.08$		.04*		
	World backward			NR	$\beta = -0.00$		.89		
	Trails B			NR	$\beta = 2.23$		.18		
	Category fluency			NR	$\beta = 0.08$		.59		
	Blessed			NR	$\beta = 0.05$		.16		
Skogen et al. (2013)	MMSE	$B = -0.03$	0.09	ns	$B = -0.03$	0.09	ns	**B:** Sex **AD:** Age	Medium
	Digit Symbol	$B = -0.12$	0.44	ns	$B = -0.14$	0.45	ns		
	Kendrick Object Learning	$B = -0.24$	0.79	ns	$B = 0.24$	0.78	ns		
	COWAT	$B = 0.85$	0.55	ns	$B = 0.91$	0.55	ns		
	Trail Making A	$B = 2.44$	2.94	ns	$B = 2.01$	2.97	ns		
	Block Design	$B = -0.23$	0.21	ns	$B = -0.26$	0.21	ns		
	Composite score	$B = 0.01$	0.10	ns	$B = 0.02$	0.10	ns		
Muller et al. (2014)	Memory			NR	$\beta = -.012$	0.22	ns	**B:** Sex **AD:** Age, education	Medium
	Processing speed			NR	$\beta = .001$	0.02	ns		
	Executive function			NR	$\beta = .010$	0.02	ns		
Shenkin et al. (2009)	RSPM	$r = .15$	ns	ns	$r = .08$	ns	ns	**B:** GA, sex, parity, social class **M:** Age	Low
	Moray House Test No. 12	$r = .15$	ns	ns	$r = .10$	ns	ns		
	Verbal Fluency	$r = .08$	ns	ns	$r = .03$	ns	ns		
	Logical Memory	$r = .09$	ns	ns	$r = .04$	ns	ns		
	g (general intelligence)	$r = .15$	ns	ns	$r = .12$	.27	.27		
	NART	$r = .10$	ns	ns	$r = .15$	.19	.19		
	g corrected for NART				$r = .05$	.63	.63		

[a] Possibly adjusted for education, as this is mentioned in-text alongside the list of confounding variables once, but not listed with them any other times.
[b] The study reported on two different age groups and has been split accordingly

Note: * indicates a significant effect was identified, + indicates that significance was reported without a p value for the linear trend, † indicates results were presented categorically; where this is not graphical, results are reported in Supplements 4 and 5 for unadjusted and adjusted conditions respectively.
Confounders are indicated relating to the participant at birth (B), the participant in childhood (CH), the participant in adulthood (AD), the mother at birth (M), the father at birth (F) or both parents at birth (P).
Abbreviations: AH4, Alice Heim Test (fourth version); Alc, alcohol intake; CLM, cholesterol lowering medication; COWAT, Controlled Oral Word Association Test; Db, diabetes; Ed, education; FC, field centre; GA, gestational age; Hdb, history of diabetes; HC, head circumference; HDL, high density liproprotein; Hhd, history of heart disease; Hst, history of stroke; HS, self-reported health status; LDL, low density liproprotein; LTM, long-term memory; MMSE, Mini-mental state examination; NART; National Adult Reading Test; NR, not recorded; RSPM; Raven's Standard Progressive Matrices; SI, sport index; STM, short-term memory; Trails-B, Trail-Making Test B; VE, vital exhaustion.

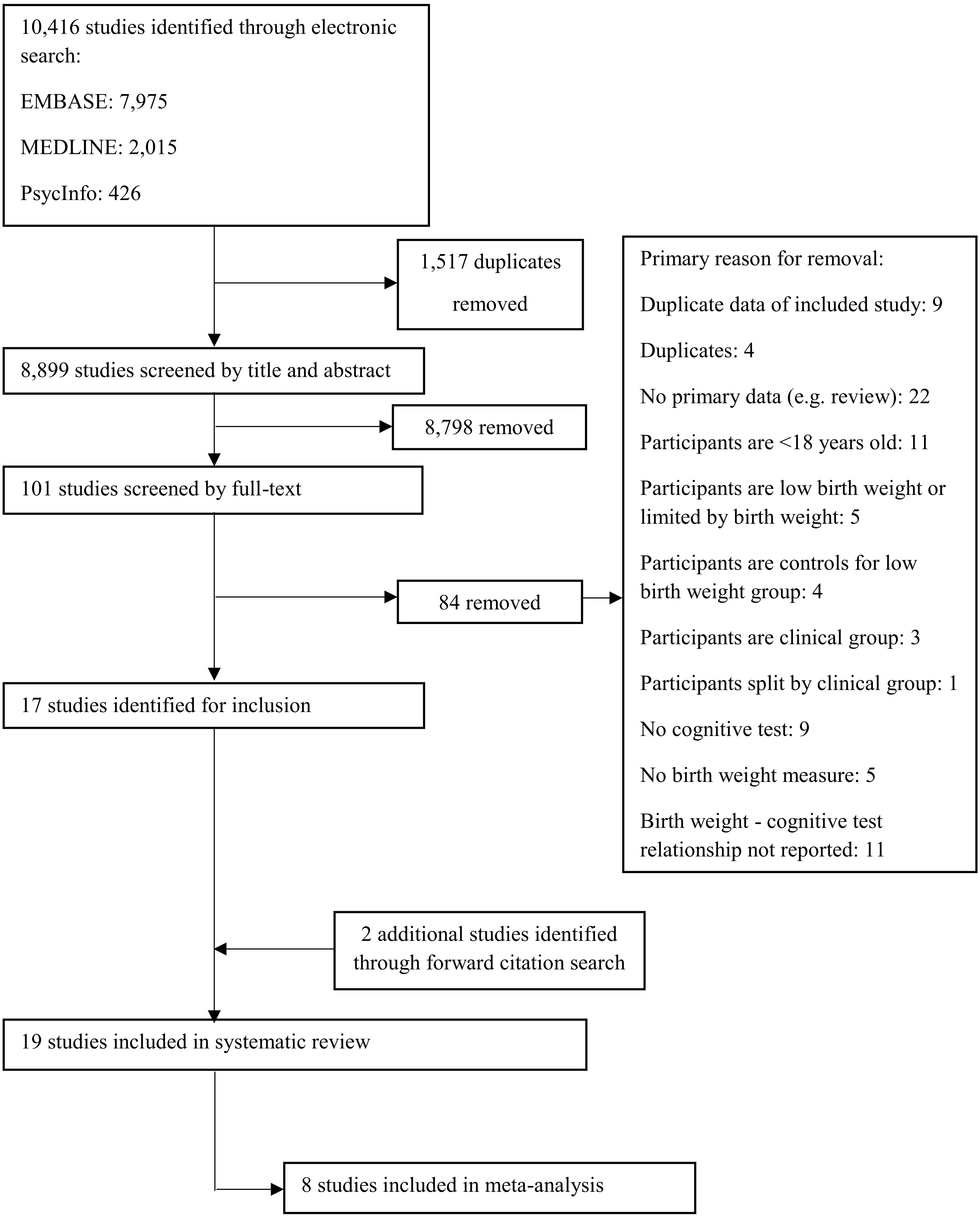

Figure 1. Study selection flow chart

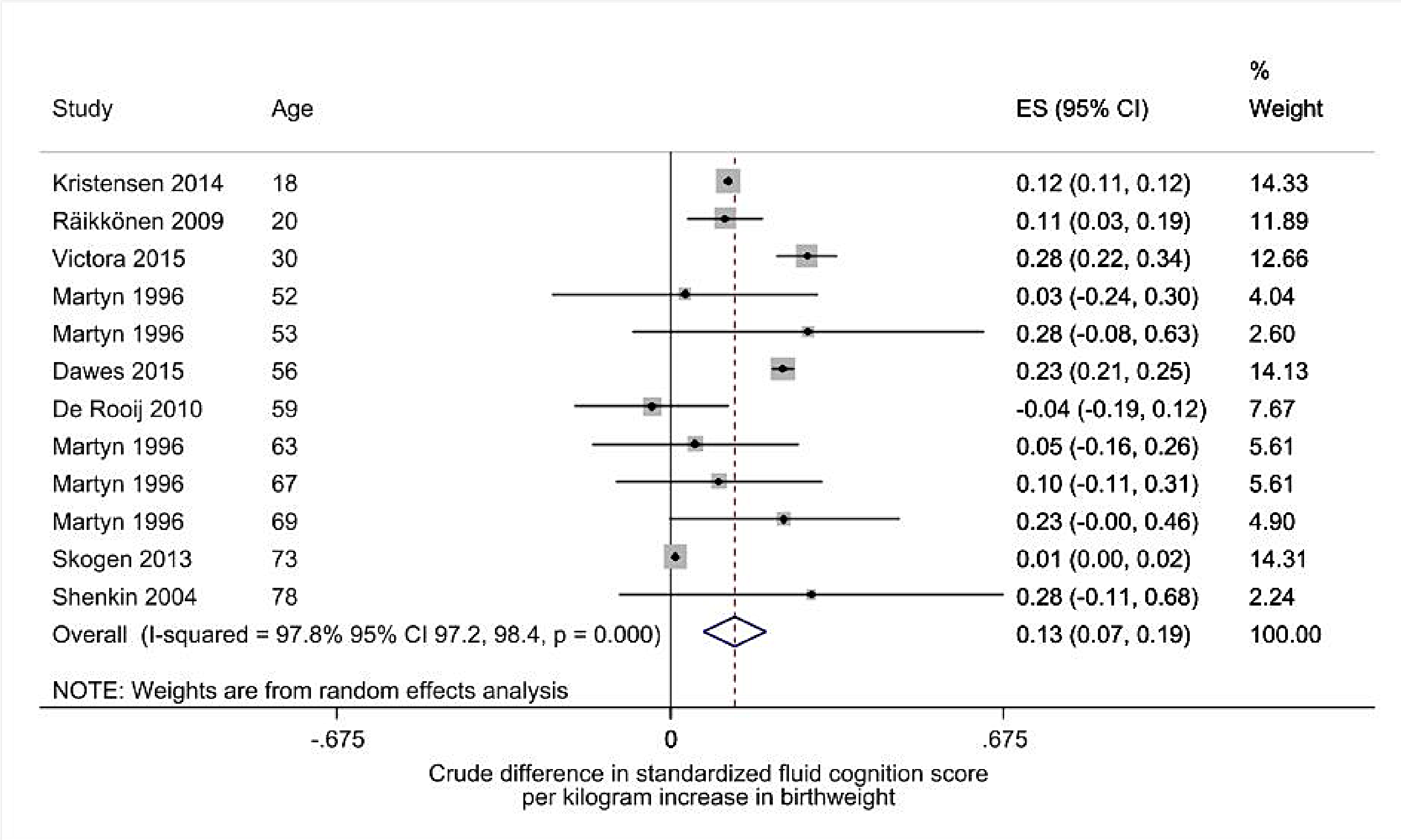

Figure 2. Forest plot of the effect size for birth weight and cognitive ability.

Note Martyn (1996) provided data for five separate cohorts, each shown separately

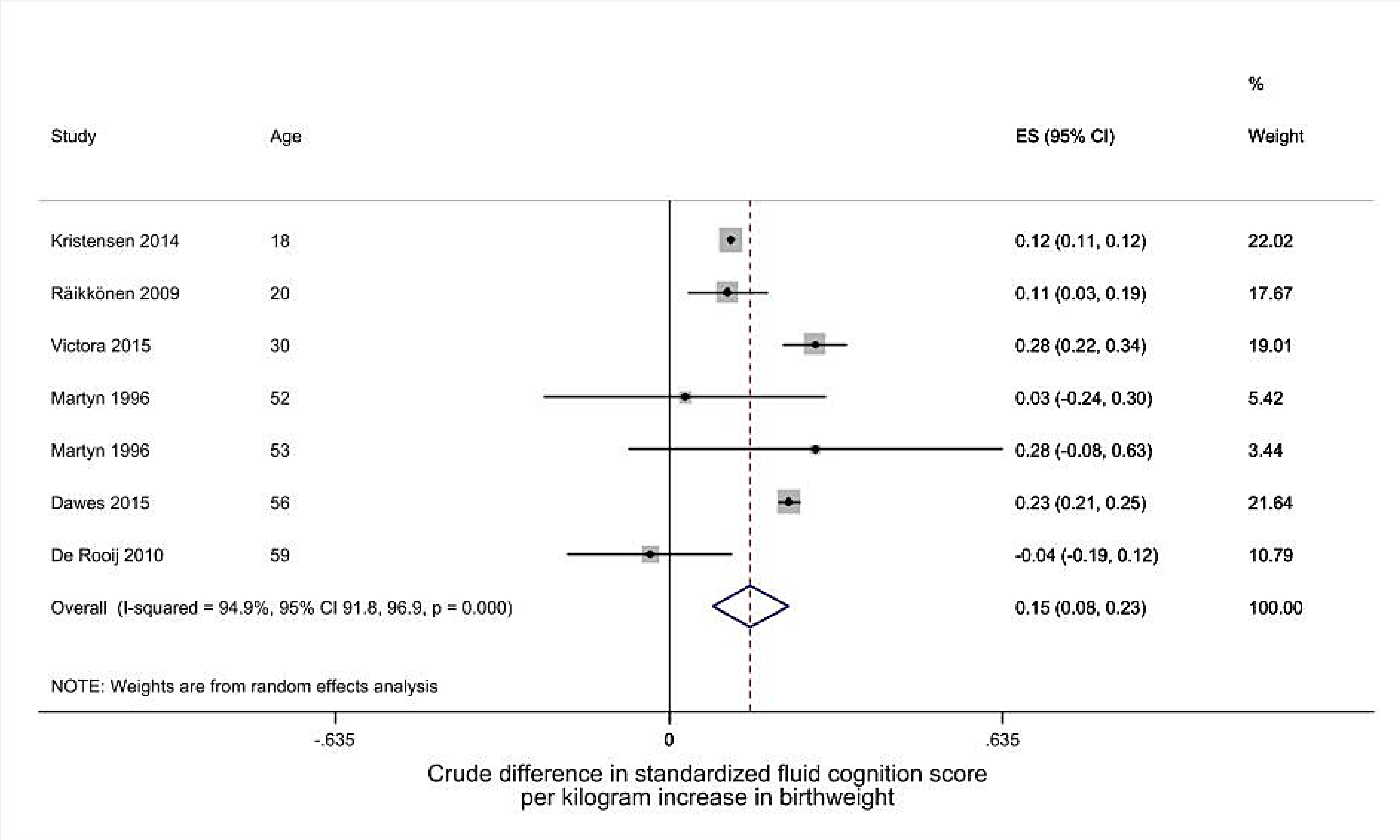

Figure 3. Forest plot of the effect size for birth weight and cognitive ability in studies with mean participant age < 60 years.

Note Martyn (1996) provided data for two separate cohorts, each shown separately

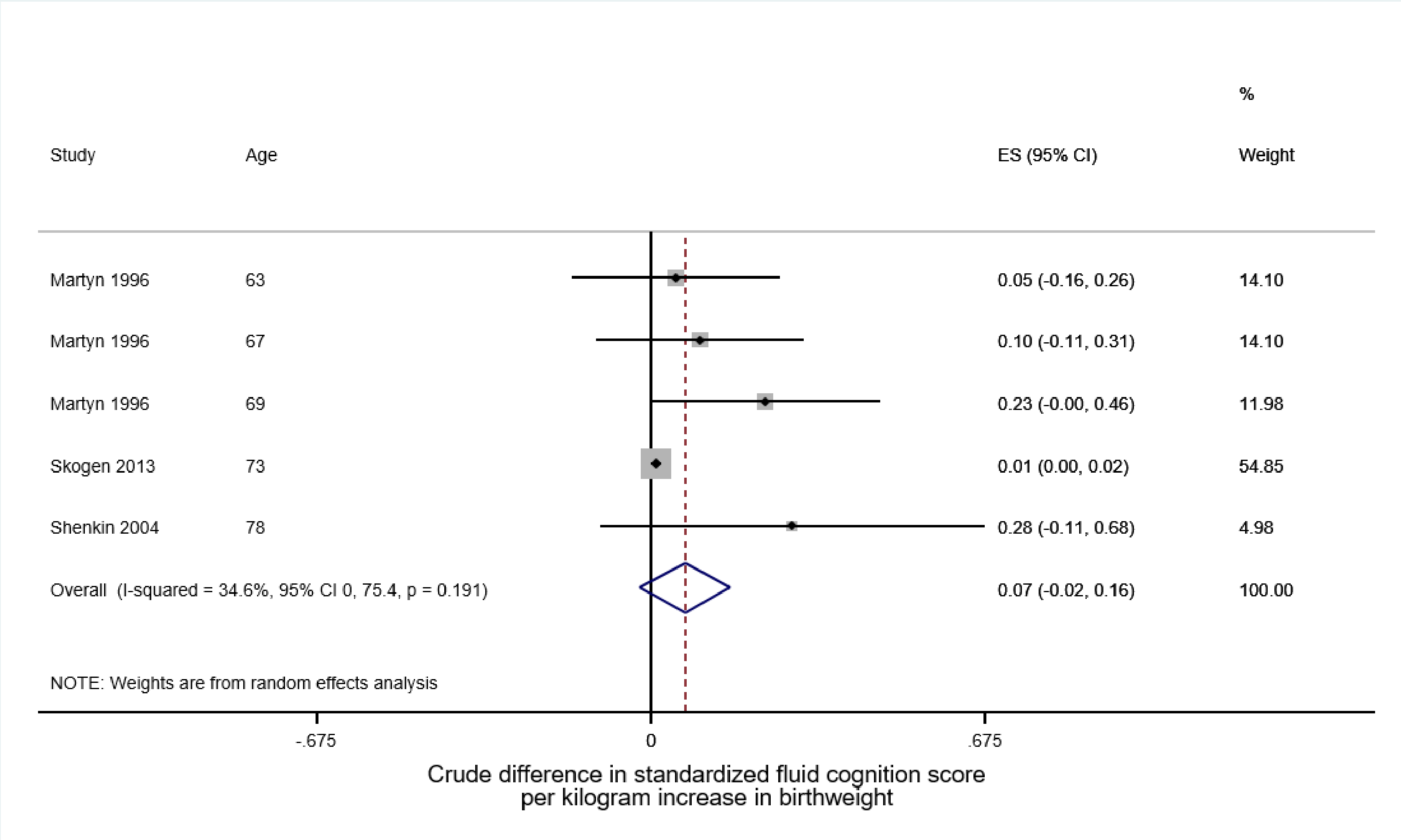

Figure 4. Forest plot of the effect size for birth weight and cognitive ability in studies with mean participant age ≥ 60 years

Note Martyn (1996) provided data for three separate cohorts, each shown separately